

## Raw Sequence Listing Error Summary

### ERROR DETECTED      SUGGESTED CORRECTION

SERIAL NUMBER: 091849,980A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☒ **Wrapped Nucleics**  
     **Wrapped Aminos**      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
  
- 2 ☐ **Invalid Line Length**      The rules require that a line not exceed 72 characters in length. This includes white spaces.
  
- 3 ☐ **Misaligned Amino**  
     **Numbering**      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
  
- 4 ☐ **Non-ASCII**      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
  
- 5 ☐ **Variable Length**      Sequence(s) \_\_\_\_\_ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
  
- 6 ☐ **PatentIn 2.0**  
     **"bug"**      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) \_\_\_\_\_. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
  
- 7 ☐ **Skipped Sequences**  
     **(OLD RULES)**      Sequence(s) \_\_\_\_\_ missing. If intentional, please insert the following lines for each skipped sequence:  
                                  (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                                  (i)        SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
                                  (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                                  This sequence is intentionally skipped  
  
                                  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
  
- 8 ☐ **Skipped Sequences**  
     **(NEW RULES)**      Sequence(s) \_\_\_\_\_ missing. If intentional, please insert the following lines for each skipped sequence.  
                                  <210> sequence id number  
                                  <400> sequence id number  
                                  000
  
- 9 ☐ **Use of n's or Xaa's**  
     **(NEW RULES)**      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
                                  Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
                                  In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
  
- 10 ☐ **Invalid <213>**  
     **Response**      Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
  
- 11 ☐ **Use of <220>**      Sequence(s) \_\_\_\_\_ missing the <220> "Feature" and associated numeric identifiers and responses.  
                                  Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
                                  (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
  
- 12 ☐ **PatentIn 2.0**  
     **"bug"**      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/849,980A

DATE: 11/10/2001  
TIME: 01:51:11

INPUT SET: S36663.raw

This Raw Listing contains the General  
Information Section and those Sequences  
containing ERRORS.

Does Not Comply  
Corrected Diskette Needed  
Errors on pp. 2-

SEQUENCE LISTING

1  
2  
3 (1) General Information  
4  
5 (i) APPLICANT: SANTEN PHARMACEUTICAL CO., LTD.  
6 (ii) TITLE OF INVENTION: Novel Polypeptide Having Water Channel  
7 Activity and DNA sequence  
8 (iii) NUMBER OF SEQUENCES: 2  
9 (iv) CORRESPONDENCE ADDRESS:  
10 (A) ADDRESSEE: SANTEN PHARMACEUTICAL CO., LTD.  
11 (B) STREET: 9 19 Shimoshinjo 3-chome Higashiyodogawa-Ku  
12 (C) CITY: Osaka  
13 (D) STATE: Osaka  
14 (E) COUNTRY: JAPAN  
15 (F) ZIP: 533-0021  
16 (v) COMPUTER READABLE FORM:  
17 (A) MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB, storage  
18 (B) COMPUTER: IBM PS/2 or compatibles  
19 (C) OPERATING SYSTEM: WINDOWS 95/97  
20 (D) SOFTWARE: Microsoft Word 97  
21 (vi) CURRENT APPLICATION DATE:  
22 (A) APPLICATION NUMBER: 09/381,810  
23 (B) FILING DATE: 19-OCT-1999  
24 (C) CLASSIFICATION: 435  
25 (vii) PRIOR APPLICATION DATA  
26 (A) APPLICATION NUMBER: PCT/JP98/01371  
27 (B) FILING DATE: 27-MAR-1998  
28 (C) APPLICATION NUMBER: JP 09-094845  
29 (D) FILING DATE: 28-MAR-1997  
30 (viii) ATTORNEY/AGENT INFORMATION:  
31 (A) NAME: Burton A. Amernick  
32 (B) REGISTRATION NUMBER: 24852  
33 (C) REFERENCE/DOCKET NUMBER: 1581/00156  
34 (ix) TELECOMMUNICATION INFORMATION:  
35 (A) TELEPHONE: (202)331-7111  
36 (B) FAX: (202)293-6229  
37

ERRORED SEQUENCES FOLLOW:

38 (2) INFORMATION FOR SEQ ID NO: 1:  
39

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/849,980A

 DATE: 11/10/2001  
 TIME: 01:51:12

INPUT SET: S36663.raw

40 (i) SEQUENCE CHARACTERISTICS:  
 --> 41 (A) LENGTH (F342) amino acids  
 42 (B) TYPE: amino acid  
 --> 43 (D) TOPOLOGY (F) linear  
 44 (ii) MOLECULE TYPE: F peptide  
 45  
 --> 46 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 47  
 48 Met Val Gln Ala Ser Gly His Arg Arg Ser Thr Arg Gly Ser Lys Met  
 49 5 10 15  
 50 Val Ser Trp Ser Val Ile Ala Lys Ile Gln Glu Ile Leu Gln Arg Lys  
 51 20 25 30  
 52 Met Val Arg Glu Phe Leu Ala Glu Phe Met Ser Thr Tyr Val Met Met  
 53 35 40 45  
 54 Val Phe Gly Leu Gly Ser Val Ala His Met Val Leu Asn Lys Lys Tyr  
 55 50 55 60  
 56 Gly Ser Tyr Leu Gly Val Asn Leu Gly Phe Gly Phe Gly Val Thr Met  
 57 65 70 75 80  
 58 Gly Val His Val Ala Gly Arg Ile Ser Gly Ala His Met Asn Ala Ala  
 59 85 90 95  
 60 Val Thr Phe Ala Asn Cys Ala Leu Gly Arg Val Pro Trp Arg Lys Phe  
 61 100 105 110  
 62 Pro Val Tyr Val Leu Gly Gln Phe Leu Gly Ser Phe Leu Ala Ala Ala  
 63 115 120 125  
 64 Thr Ile Tyr Ser Leu Phe Tyr Thr Ala Ile Leu His Phe Ser Gly Gly  
 65 130 135 140  
 66 Gln Leu Met Val Thr Gly Pro Val Ala Thr Ala Gly Ile Phe Ala Thr  
 67 145 150 155 160  
 68 Tyr Leu Pro Asp His Met Thr Leu Trp Arg Gly Phe Leu Asn Glu Ala  
 69 165 170 175  
 70 Trp Leu Thr Gly Met Leu Gln Leu Cys Leu Phe Ala Thr Thr Asp Gln  
 71 180 185 190  
 72 Glu Asn Asn Pro Ala Leu Pro Gly Thr Glu Ala Leu Val Ile Gly Ile  
 73 195 200 205  
 74 Leu Val Val Ile Ile Gly Val Ser Leu Gly Met Asn Thr Gly Tyr Ala  
 75 210 215 220  
 76 Ile Asn Pro Ser Arg Asp Leu Pro Pro Arg Ile Phe Thr Phe Ile Ala  
 77 225 230 235 240  
 78 Gly Trp Gly Lys Gln Val Phe Ser Asn Gly Glu Asn Trp Trp Trp Val  
 79 245 250 255  
 80 Pro Val Val Ala Pro Leu Leu Gly Ala Tyr Leu Gly Gly Ile Ile Tyr  
 81 260 265 270  
 82 Leu Val Phe Ile Gly Ser Thr Ile Pro Arg Glu Pro Leu Lys Leu Glu  
 83 275 280 285  
 84 Asp Ser Val Ala Tyr Glu Asp His Gly Ile Thr Val Leu Pro Lys Met  
 85 290 295 300  
 86 Gly Ser His Glu Pro Thr Ile Ser Pro Leu Thr Pro Val Ser Val Ser  
 87 305 310 315 320  
 88 Pro Ala Asn Arg Ser Ser Val His Pro Ala Pro Pro Leu His Glu Ser  
 89 325 330 335  
 90 Met Ala Leu Glu His Phe  
 91 340  
 92

 invalid:  
 cannot use "F"

# RAW SEQUENCE LISTING

## PATENT APPLICATION US/09/849,980A

DATE: 11/10/2001  
TIME: 01:51:12

INPUT SET: S36663.raw

```

93 (2) INFORMATION FOR SEQ ID NO:2:
94 (i) SEQUENCE CHARACTERISTICS:
--> 95 (A) LENGTH: F1258 base pairs
96 (B) TYPE: nucleic acid
97 (C) STRANDEDNESS: double
--> 98 (D) TOPOLOGY: F linear
99 (ii) MOLECULE TYPE: cDNA to mRNA
100 (vi) ORIGINAL SOURCE:
101 (A) ORGANISM: Homo sapiens
102 (B) TISSUE TYPE: fat tissue
103 (ix) FEATURE:
104 (A) NAME/KEY: exon
105 (B) LOCATION: F173..1198
106 (C) IDENTIFICATION METHOD: by experiment
107 (C) IDENTIFICATION METHOD: by experiment
--> 108 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
109
110 GGCTCTGGAC TGGGGACACA GGGATAGCTG AGCCCCAGCT
--> 111 GGGGGTGGAA GCTGAGCCAG 60
112 GGACAGTCAC GGAGGAACAA GATCAAGATG CGCTGTAAC
--> 113 GAGAAGCCCC CAAGGCGGAG 120
114 GCTGAGAATC AGAGACATTT CAGCAGACAT CTACAAATCT
--> 115 GAAAGACAAA AC ATG GTT 178
116
117
118 CAA GCA TCC GGG CAC AGG CGG TCC ACC CGT GGC TCC AAA
--> 119 ATG GTC TCC 226
120 Gln Ala Ser Gly His Arg Arg Ser Thr Arg Gly Ser Lys Met Val Ser
121 5 10 15
122 TGG TCC GTG ATA GCA AAG ATC CAG GAA ATA CTG CAG AGG
--> 123 AAG ATG GTG 274
124 Trp Ser Val Ile Ala Lys Ile Gln Glu Ile Leu Gln Arg Lys Met Val
125 20 25 30
126 CGA GAG TTC CTG GCC GAG TTC ATG AGC ACA TAT GTC ATG
--> 127 ATG GTA TTC 322
128 Arg Glu Phe Leu Ala Glu Phe Met Ser Thr Tyr Val Met Met Val Phe
129 35 40 45 50
130 GGC CTT GGT TCC GTG GCC CAT ATG GTT CTA AAT AAA AAA
--> 131 TAT GGG AGC 370
132 Gly Leu Gly Ser Val Ala His Met Val Leu Asn Lys Lys Tyr Gly Ser
133 55 60 65
134 TAC CTT GGT GTC AAC TTG GGT TTT GGC TTC GGA GTC ACC
--> 135 ATG GGA GTG 418
136 Tyr Leu Gly Val Asn Leu Gly Phe Gly Phe Gly Val Thr Met Gly Val
137 70 75 80
138 CAC GTG GCA GGC CGC ATC TCT GGA GCC CAC ATG AAC GCA
--> 139 GCT GTG ACC 466
140 His Val Ala Gly Arg Ile Ser Gly Ala His Met Asn Ala Ala Val Thr
141 85 90 95
142 TTT GCT AAC TGT GCG CTG GGC CGC GTG CCC TGG AGG AAG
--> 143 TTT CCG GTC 514
144 Phe Ala Asn Cys Ala Leu Gly Arg Val Pro Trp Arg Lys Phe Pro Val

```

text -  
wrapped - see error summary  
sheet, item 1

Met Val  
1

Same as  
above

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/849,980A

DATE: 11/10/2001  
TIME: 01:51:13

INPUT SET: S36663.raw

```

145      100      105      110
146 TAT GTG CTG GGG CAG TTC CTG GGC TCC TTC CTG GCG GCT
--> 147 GCC ACC ATC      562
148 Tyr Val Leu Gly Gln Phe Leu Gly Ser Phe Leu Ala Ala Ala Thr Ile
149      115      120      125      130
150 TAC AGT CTC TTC TAC ACG GCC ATT CTC CAC TTT TCG GGT
--> 151 GGA CAG CTG      610
152 Tyr Ser Leu Phe Tyr Thr Ala Ile Leu His Phe Ser Gly Gly Gln Leu
153      135      140      145
154 ATG GTG ACC GGT CCC GTC GCT ACA GCT GGC ATT TTT GCC
--> 155 ACC TAC CTT      658
156 Met Val Thr Gly Pro Val Ala Thr Ala Gly Ile Phe Ala Thr Tyr Leu
157      150      155      160
158 CCT GAT CAC ATG ACA TTG TGG CGG GGC TTC CTG AAT GAG
--> 159 GCG TGG CTG      706
160 Pro Asp His Met Thr Leu Trp Arg Gly Phe Leu Asn Glu Ala Trp Leu
161      165      170      175
162 ACC GGG ATG CTC CAG CTG TGT CTC TTC GCC ATC ACG GAC
--> 163 CAG GAG AAC      754
164 Thr Gly Met Leu Gln Leu Cys Leu Phe Ala Thr Thr Asp Gln Glu Asn
165      180      185      190
166 AAC CCA GCA CTG CCA GGA ACA GAG GCG CTG GTG ATA
--> 167 GGC ATC CTC GTG      802
168 Asn Pro Ala Leu Pro Gly Thr Glu Ala Leu Val Ile Gly Ile Leu Val
169      195      200      205      210
170 GTC ATC ATC GGG GTG TCC CTT GGC ATG AAC ACA GGA TAT
--> 171 GCC ATC AAC      850
172 Val Ile Ile Gly Val Ser Leu Gly Met Asn Thr Gly Tyr Ala Ile Asn
173      215      220      225
174 CCG TCC CGG GAC CTG CCC CCC CGC ATC TTC ACC TTC ATT
--> 175 GCT GGT TGG      898
176 Pro Ser Arg Asp Leu Pro Pro Arg Ile Phe Thr Phe Ile Ala Gly Trp
177      230      235      240
178 GGC AAA CAG GTC TTC AGC AAT GGG GAG AAC TGG TGG
--> 179 TGG GTG CCA GTG      946
180 Gly Lys Gln Val Phe Ser Asn Gly Glu Asn Trp Trp Trp Val Pro Val
181      245      250      255
182 GTG GCA CCA CTT CTG GGT GCC TAT CTA GGT GGC ATC ATC
--> 183 TAC CTG GTC      994
184 Val Ala Pro Leu Leu Gly Ala Tyr Leu Gly Gly Ile Ile Tyr Leu Val
185      260      265      270
186 TTC ATT GGC TCC ACC ATC CCA CGG GAG CCC CTG AAA TTG
--> 187 GAG GAT TCT      1042
188 Phe Ile Gly Ser Thr Ile Pro Arg Glu Pro Leu Lys Leu Glu Asp Ser
189      275      280      285      290
190 GTG GCG TAT GAA GAC CAC GGG ATA ACC GTA TTG CCC AAG
--> 191 ATG GGA TCT      1090
192 Val Ala Tyr Glu Asp His Gly Ile Thr Val Leu Pro Lys Met Gly Ser
193      295      300      305
194 CAT GAA CCC ACG ATC TCT CCC CTC ACC CCC GTC TCT GTG
--> 195 AGC CCT GCC      1138
196 His Glu Pro Thr Ile Ser Pro Leu Thr Pro Val Ser Val Ser Pro Ala
197      310      315      320

```

Same

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/849,980ADATE: 11/10/2001  
TIME: 01:51:13

INPUT SET: S36663.raw

--> 198 AAC AGA TCT TCA GTC CAC CCT GCC CCA CCC TTA CAT GAA  
199 TCC ATG GCC 1186  
200 Asn Arg Ser Ser Val His Pro Ala Pro Pro Leu His Glu Ser Met Ala  
201 325 330 335  
202 CTA GAG CAC TTC TAAGCAGAGA TTATTGTGA TCCCATCCAT  
--> 203 TCCCAATAA 1238  
204 Leu Glu His Phe  
205 340  
206 AGCAAGGCTT GTCCGACAAA  
207

1258

# SEQUENCE VERIFICATION REPORT

## PATENT APPLICATION US/09/849,980A

DATE: 11/10/2001  
TIME: 01:51:13

INPUT SET: S36663.raw

Line	Error	Original Text
22	Wrong application Serial Number	(A) APPLICATION NUMBER: 09/381,810
24	Wrong Classification	(C) CLASSIFICATION: 435
41	Length must be an Integer	(A) LENGTHF342 amino acids
41	Entered (0) and Calc. Seq. Length (342) differ	(A) LENGTHF342 amino acids
43	Wrong or Missing Sequence Topology	(D) TOPOLOGYF linear
46	Wrong or Missing Sequence Topology	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
95	Length must be an Integer	(A) LENGTHF1258 base pairs
95	Entered (0) and Calc. Seq. Length (283) differ	(A) LENGTHF1258 base pairs
98	Wrong or Missing Sequence Topology	(D) TOPOLOGYF linear
108	Wrong or Missing Sequence Topology	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
111	# of Sequences for line conflicts w/ running total	GGGGGTGGAA GCTGAGCCAG 60
113	# of Sequences for line conflicts w/ running total	GAGAAGCCCC CAAGGCGGAG 120
115	# of Sequences for line conflicts w/ running total	GAAAGACAAA AC ATG GTT 178
119	# of Sequences for line conflicts w/ running total	ATG GTC TCC 226
123	# of Sequences for line conflicts w/ running total	AAG ATG GTG 274
127	# of Sequences for line conflicts w/ running total	ATG GTA TTC 322
131	# of Sequences for line conflicts w/ running total	TAT GGG AGC 370
135	# of Sequences for line conflicts w/ running total	ATG GGA GTG 418
139	# of Sequences for line conflicts w/ running total	GCT GTG ACC 466
143	# of Sequences for line conflicts w/ running total	TTT CCG GTC 514
147	# of Sequences for line conflicts w/ running total	GCC ACC ATC 562
151	# of Sequences for line conflicts w/ running total	GGA CAG CTG 610
155	# of Sequences for line conflicts w/ running total	ACC TAC CTT 658
159	# of Sequences for line conflicts w/ running total	GCG TGG CTG 706
163	# of Sequences for line conflicts w/ running total	CAG GAG AAC 754
167	# of Sequences for line conflicts w/ running total	GGC ATC CTC GTG 802
171	# of Sequences for line conflicts w/ running total	GCC ATC AAC 850
175	# of Sequences for line conflicts w/ running total	GCT GGT TGG 898
179	# of Sequences for line conflicts w/ running total	TGG GTG CCA GTG 946
183	# of Sequences for line conflicts w/ running total	TAC CTG GTC 994
187	# of Sequences for line conflicts w/ running total	GAG GAT TCT 1042
191	# of Sequences for line conflicts w/ running total	ATG GGA TCT 1090
195	# of Sequences for line conflicts w/ running total	AGC CCT GCC 1138
199	# of Sequences for line conflicts w/ running total	TCC ATG GCC 1186
203	# of Sequences for line conflicts w/ running total	TCCCCAATAA 1238

**SEQUENCE CORRECTION REPORT**  
**PATENT APPLICATION US/09/849,980A**DATE: 11/10/2001  
TIME: 01:51:13**INPUT SET: S36663.raw**

Line	Original Text	Corrected Text
3	(1) General Information	(1) GENERAL INFORMATION:
21	(vi) CURRENT APPLICATION DATE:	(vi) CURRENT APPLICATION DATA:
25	(vii) PRIOR APPLICATION DATA	(vii) PRIOR APPLICATION DATA:
29	(D) FILILNG DATE: 28-MAR-1997	(D) FILING DATE: 28-MAR-1997
36	(B) FAX: (202)293-6229	(B) TELEFAX: (202)293-6229
44	(ii) MOLECULE TYPEF peptide	(ii) MOLECULE TYPE:F peptide
93	(2) INFORMATION FOR SEQ ID NoF2	(2) INFORMATION FOR SEQ ID NO:2:
105	(B) LOCATIONF173..1198	(B) LOCATION:F173..1198